

Data Science Integration

Change Log

Version	Date	Author	Comments
0.1	6/14/18	Kirk Alford	Initial Draft

Overview

The data science team has and will continue to develop new model leading to new concepts needed to be stored within the CDP to leverage these concepts in future campaigns. Integration between the Data Pipelines and Data Science is needed to get the data generated from Data Science available in the CDP.

JIRA Epics

Key	Summary	Due	Status
EDH-1972	Data Science Integration		OPEN
DSCI-25	CDP Integration		BACKLOG

[2 issues](#)

Scope

- The integration between data science and the CDP will allow for the seamless introduction of new concepts based on data science output. Once complete, minimal, if any effort will be required from the CDP team to add new concepts to the CDP.
- The integration should provide a mechanism to determine a derived value, preferably as an enrichment job.
 - This should account for existing persons as when a new concept is introduced as well as with ongoing updates and new persons being added to the datastore.
 - Example, patient value max is released. The application should determine the patient value max for any existing persons and then for any new persons added with a future data update.
- Data updates to a customer's database should also be made available in the master model matrix to increase the size of the data set used for data science.

Diagrams

High level workflow diagrams.

Out of Scope

This section should be used to list functionality not included in the feature/product. This section is often populated when an item that was previously in scope is changed to out of scope.

Risks

Optional section. If there are significant risks or challenges to implementing the feature (e.g. dependency on another vendor), please list them here.

JIRA User Stories

Here is where we will add a JIRA widget that will show all related user stories to the feature. You will need to edit the criteria to get the correct stories to show up. You should be able to just add the epic id's (*epic_link*) to get the correct stories to display.

Key	Status	Summary
EDH-2478	NEW	Automate the scoring for data models in the Tanner index
EDH-2477	NEW	Get scores for each person record from Data Science

2 issues